Yannick Deville
Sharon Gannot
Russell Mason
Mark D. Plumbley
Dominic Ward (Eds.)

LNCS 10891

Latent Variable Analysis and Signal Separation

14th International Conference, LVA/ICA 2018
Guildford, UK, July 2–5, 2018
Proceedings

Springer

Applications of Polynomial Common Factor Computation in Signal Processing

Ivan Markovsky[1(✉)], Antonio Fazzi[2], and Nicola Guglielmi[2]

[1] Department ELEC, Vrije Universiteit Brussel (VUB), Pleinlaan 2, Building K, 1050 Brussels, Belgium
imarkovs@vub.ac.be

[2] Gran Sasso Science Institute (GSSI), 67010 L' Aquila, Italy
antonio.fazzi@gssi.it, guglielm@univaq.it

Abstract. We consider the problem of computing the greatest common divisor of a set of univariate polynomials and present applications of this problem in system theory and signal processing. One application is blind system identification: given the responses of a system to unknown inputs, find the system. Assuming that the unknown system is finite impulse response and at least two experiments are done with inputs that have finite support and their Z-transforms have no common factors, the impulse response of the system can be computed up to a scaling factor as the greatest common divisor of the Z-transforms of the outputs. Other applications of greatest common divisor problem in system theory and signal processing are finding the distance of a system to the set of uncontrollable systems and common dynamics estimation in a multi-channel sum-of-exponentials model.

Keywords: Blind system identification
Sum-of-exponentials modeling · Distance to uncontrollability
Approximate common factor · Low-rank approximation

1 Introduction

Finding the *greatest common divisor* of a set of univariate polynomials is a classic problem in algebra, which is still an active research topic. Numerically it is an ill-conditioned problem: small perturbations in the input data (the polynomials' coefficients) may result in large changes in the solution (the greatest common divisor coefficients). This requires computing an *approximate* common divisor.

There are two different formulations of the approximate common divisor problem. In the first one, the degree of the common divisor is a priori specified and the smallest perturbation on the polynomial coefficients that leads to polynomials with common divisor of such a degree is sought [6,9,13,15,18]. In the second formulation, refered to as *ϵ-common divisor*, the size of the maximum perturbation is given and perturbed polynomials with maximal degree common divisor is sought [1,14,17]. The two problems are dual to each other. In fact,

Y. Deville et al. (Eds.): LVA/ICA 2018, LNCS 10891, pp. 99–106, 2018.
https://doi.org/10.1007/978-3-319-93764-9_10

they are two different scalarizations of the biobjective problem where the size of the perturbation is minimized while the degree of the perturbed polynomials common divisor is maximized. The two formulations trace the same Pareto optimal trade-off curve.

The approximate greatest common divisor problem is a non-convex optimization problem, for which there are no efficient global solution methods. The existing methods can be classified as local optimization methods and convex relaxations. Local optimization methods require an initial approximation and are in general computationally more expensive than the relaxation methods, however, the local optimization methods explicitly optimize the desired criterion (size of the coefficient perturbation), which ensures that they produce at least as good result as a relaxation method, provided the solution of the relaxation method is used as an initial approximation for the local optimization method. For a recent overview of computational approaches, we refer the reader to [15].

Applications of the greatest common divisor in systems, control, and signal processing, however, are surprisingly missing from the broad literature on the theoretical and computational aspects of the problem. We present here applications that are directly solvable by a greatest common divisor computation. Subsequently existing greatest common divisor methods, algorithms and software can be used in the applications. Vice versa, methods, algorithms and software developed for the applications can be viewed as and used for greatest common divisor computation.

In this paper, we consider the following approximate common factor computation problem: given polynomials $p^1, \ldots, p^N$ and a natural number $\mathtt{d}$,

$$\begin{aligned} &\text{minimize} \quad \text{over } \widehat{p}^1, \ldots, \widehat{p}^N \quad \sum_{i=1}^{N} \|p^i - \widehat{p}^i\|_2^2 \\ &\text{subject to} \quad \deg\big(\gcd(\widehat{p}^1, \ldots, \widehat{p}^N)\big) \geq \mathtt{d}. \end{aligned} \tag{1}$$

($\gcd(p^1, \ldots, p^N)$ is the greatest common divisor of the polynomials $p^1(z), \ldots, p^N(z)$.) Sect. 2 shows application of (1) for blind finite impulse response system identification. Section 3 shows application of (1) for computing the distance of a given linear time-invariant system to the set of uncontrollable systems. Section 4 shows application of (1) for estimation of common dynamics across multiple channels of an autonomous linear time-invariant system.

2 Blind Finite Impulse Response System Identification

The identification problem considered in this section is defined as follows.

Problem 1 (Blind finite impulse response system identification). Given output observations $y^1, \ldots, y^N$ of a *finite impulse response* system, generated by unknown signals $u^1, \ldots, u^N$, find the impulse response h of the system.

The Z-transform of a finite duration time-domain signal y^i is a polynomial $y^i(z)$. (We use the argument z, as in $y^i(z)$, to indicate that the signal is in the Z-domain).

Theorem 1. *Assuming that at least $N = 2$ responses $y^1, \ldots, y^N$ of a finite impulse response system are given,*

1. *the inputs $u^1, \ldots, u^N$ have finite support, and*
2. $\gcd\left(u^1(z), \ldots, u^N(z)\right) = 1$,

the impulse response h of the system is up to a scaling factor $\alpha \in \mathbb{R}$ the greatest common factor of $y^1(z), \ldots, y^N(z)$,

$$h(z) = \alpha \gcd\left(y^1(z), \ldots, y^N(z)\right).$$

Proof. Let $\star$ be the convolution operator. We have,

$$y^i = h \star u^i, \quad \text{for } i = 1, \ldots, N.$$

Since the system is finite impulse response $h(z) := Z(h)$ is a polynomial. Under assumption 1, $u^i(z) := Z(u^i)$ are also polynomials. Therefore, $y^i(z) := Z(y^i)$

$$y^i(z) = h(z)u^i(z), \quad \text{for } i = 1, \ldots, N \tag{2}$$

are polynomials. It follows from (2) that $h(z)$ is a common factor of $y^1, \ldots, y^N$. By assumption 2, $h(z)$ is the greatest common factor of $y^1, \ldots, y^N$. □

With noisy data

$$y_{\mathrm{d}}^i = \bar{y}^i + \tilde{y}^i, \quad \text{for } i = 1, \ldots, N$$

(the subscript index "d" stands for "data"), where $\bar{y}^i$ is the "true value" and $\tilde{y}$ is the measurement noise, $y_{\mathrm{d}}^1, \ldots, y_{\mathrm{d}}^N$ are generically co-prime, *i.e.*, they have no nontrivial common factor. Assuming that the noise $\tilde{y}$ is zero mean, white, Gaussian, the maximum-likelihood estimator of the "true impulse response" $\bar{h}$ is given by the following problem

$$\begin{aligned} &\text{minimize} \quad \text{over } \widehat{y}^1, \ldots, \widehat{y}^N, \widehat{u}^1, \ldots, \widehat{u}^N, \text{and } \widehat{h} \quad \sum_{i=1}^{N} \|y_{\mathrm{d}}^i - \widehat{y}^i\|_2^2 \\ &\text{subject to} \quad \widehat{y}^i = \widehat{h} \star \widehat{u}^i, \quad \text{for } i = 1, \ldots, N. \end{aligned} \tag{3}$$

Note that since we include $\widehat{h}$ as an optimization variable, we assume that its length (or equivalently the order $\bar{n} = \dim(\bar{h}) - 1$ of the true system) is a priori known.

Problem (3) is an approximate common factor computation problem.

Theorem 2. *Problems (3) and (1) are equivalent.*

3 Distance to Uncontrollability

Verifying whether a given linear time-invariant system is controllable involves rank computation. Arbitrary small perturbations of the system's parameters can switch the property. This issue is addressed by the notion of distance to uncontrollability, which is quantitative rather than qualitative measure of controllability. The definition of distance to uncontrollability, considered in the literature [10], is a property of the parameters A and B in a state space representation of the system. Using the notion of controllability in the behavioral setting [12] we define a representation invariant measure of distance to uncontrollability and propose an algorithm for computing it.

Consider a linear time-invariant system $\mathscr{B}$ with a state space representation

$$\mathscr{B} = \mathscr{B}_{\mathrm{i/s/o}}(A, B, C, D) := \{\, w = (u, y) \mid \sigma x = Ax + Bu,\ y = Cx + Du \,\}, \quad (4)$$

where $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$, $B \in \mathbb{R}^{\mathtt{n}\times\mathtt{m}}$, $C \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$, and $D \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}}$ are parameters of $\mathscr{B}$; and σ is the shift operator $(\sigma x)(t) = x(t+1)$ (in discrete-time) or the derivative operator $\sigma x = dx/dt$ (in continuous-time).

We adopt the behavioral setting [12], *i.e.*, a system is viewed as a set of trajectories. For a given system $\mathscr{B}$, the parameters A, B, and C of the state space representation (4) of $\mathscr{B}$ are not unique due to the fact that for any change of basis $x' = Vx$ of the state space, $\mathscr{B}(VAV^{-1}, VB, CV^{-1}, D)$ is the same model as $\mathscr{B}(A, B, C, D)$, *i.e.*,

$$\mathscr{B}_{\mathrm{i/s/o}}(A, B, C, D) = \mathscr{B}_{\mathrm{i/s/o}}(VAV^{-1}, VB, CV^{-1}, D).$$

In addition, the parameters A, B, and C are not unique due to nonminimality of the state dimension; for example

$$\mathscr{B}_{\mathrm{i/s/o}}(A, B, C, D) = \mathscr{B}\left(\begin{bmatrix} A & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \begin{bmatrix} B \\ 0 \end{bmatrix}, \begin{bmatrix} C & 0 \end{bmatrix}, D\right),$$

for any $A_{12} \in \mathbb{R}^{\mathtt{n}\times\Delta n}$, $A_{21} \in \mathbb{R}^{\Delta n\times\mathtt{n}}$, and $A_{22} \in \mathbb{R}^{\Delta\mathtt{n}\times\Delta\mathtt{n}}$.

A state space representation with parameters A and B is *state controllable* if and only if the matrix

$$\mathscr{C}(A, B) := \begin{bmatrix} A & AB & \cdots & A^{\mathtt{n}-1}B \end{bmatrix}$$

is full rank. Note that this classical notion of controllability is a property of the pair of matrices (A, B) and is not a property of a system $\mathscr{B}$ due to the nonuniqueness of a state space representation. The question of whether a given state space representation is state controllable is a rank test problem for the structured matrix $\mathscr{C}(A, B)$. A corresponding quantitative measure is the distance of $\mathscr{C}(A, B)$ to rank deficiency, *i.e.*, the smallest $(\Delta A, \Delta B)$, such that

$$\mathscr{C}(\widehat{A}, \widehat{B}) := \mathscr{C}(A, B) + \mathscr{C}(\Delta A, \Delta B)$$

is rank deficient.

Consider the set of $m \times n$ structured matrices $\mathscr{S}$ and define the distance measure

$$d_r(A) := \min_{\Delta A \in \mathscr{S}} \|\Delta A\| \quad \text{subject to} \quad A + \Delta A \text{ has rank} r,$$

where $\|\cdot\|$ is a matrix norm. With $\mathscr{S} = \mathbb{R}^{m\times n}$, $d_r(A)$ is a distance to *unstructured* rank-r matrices. In the special cases of spectral and Frobenius norms, the unstructured distance $d_r(A)$ can be computed using the singular value decomposition of A.

Motivated by the issues of computing the numerical rank of a matrix, C. Paige defined in [10] the distance to uncontrollability

$$\begin{aligned} d_{\text{unctr}}(A,B) := \text{minimize} \quad & \text{over} \widehat{A}, \widehat{B} \quad \left\| \begin{bmatrix} A & B \end{bmatrix} - \begin{bmatrix} \widehat{A} & \widehat{B} \end{bmatrix} \right\|_{\text{F}} \\ \text{subject to} \quad & (\widehat{A}, \widehat{B}) \text{ is uncontrollable.} \end{aligned}$$

This problem falls into a broader category of *distance problems* [4], such as distance to instability, distance to positive definiteness, *etc.* There is a big volume of literature devoted on the problem of computing $d_{\text{unctr}}(A,B)$, see, *e.g.*, [2,3,5,7,8]. The measure $d_{\text{unctr}}(A,B)$, however, is not invariant of the state space representation because it depends on the choice of basis. This issue is resolved in the behavioral setting, where controllability is defined as a property of the system rather then a property of a particular representation.

Definition 1 ([16]). *A time-invariant system $\mathscr{B}$ is controllable if for any two trajectories $w_p, w_f \in \mathscr{B}$, there is $\Delta t > 0$ and a trajectory $w_c \in \mathscr{B}$, such that $w_p(t) = w_c(t)$, for all $t < 0$, and $w_f(t) = w_f(t)$, for all $t \geq \Delta t$.*

Checking the controllability property in practice is done by performing a numerical test on the parameters of a specific representation of the system. For a single-input single-output linear time-invariant system with an input/output representation

$$\mathscr{B}_{\text{i/o}}(p,q) := \{ \begin{bmatrix} u \\ y \end{bmatrix} \mid p(\sigma)y = q(\sigma)u \} \tag{5}$$

is controllable if and only if the polynomials are co-prime.

Theorem 3 ([12]). *Consider the polynomials $p(z)$ and $q(z)$ and let the degree of p be higher than or equal to the degree of q. The single-input single-output system $\mathscr{B}_{i/o}(p,q)$ is controllable if and only if p and q are co-prime.*

By Theorem 3, the system $\mathscr{B}_{\text{i/o}}(p,q)$ is controllable if and only if p and q have no common factors of degree one or more.

Let $\overline{\mathscr{L}_{\text{ctrb}}}$ be the set of uncontrollable linear time-invariant systems:

$$\overline{\mathscr{L}_{\text{ctrb}}} = \{ \mathscr{B} \ : \ \mathscr{B} \text{is linear time-invariant and uncontrollable} \}$$

and consider the distance measure

$$\text{dist}\left(\mathscr{B}_{\text{i/o}}(p,q), \mathscr{B}_{\text{i/o}}(\widehat{p},\widehat{q})\right) := \left\| \begin{bmatrix} q \\ p \end{bmatrix} - \begin{bmatrix} \widehat{q} \\ \widehat{p} \end{bmatrix} \right\|_2. \tag{6}$$

The representation invariant notion of distance to uncontrollability proposed is: Given a controllable system $\mathscr{B}_{\mathrm{i/o}}(p, q)$, find:

$$d_{\mathrm{unctr}}(\mathscr{B}) := \min_{\widehat{\mathscr{B}} \in \overline{\mathscr{L}_{\mathrm{ctrb}}}} \operatorname{dist}(\mathscr{B}, \widehat{\mathscr{B}}). \tag{7}$$

We refer to $d_{\mathrm{unctr}}(\mathscr{B})$ as the *uncontrollability radius*.

Theorem 4. *Problems* (7) *and* (1) *with* $\mathtt{d} = 1$ *are equivalent.*

Proof. Follows directly from Theorem 3. □

4 Common Dynamics Estimation

The problem considered in this section is defined as follows.

Problem 2. Given a set of N scalar autonomous linear time-invariant systems $\mathscr{B}_1, \ldots, \mathscr{B}_N$, find their "common dynamics", defined as $\mathscr{B} := \mathscr{B}_1 \cap \cdots \cap \mathscr{B}_N$.

Let the systems be represented by their kernel representations $\mathscr{B}_i = \ker\big(p^i(\sigma)\big)$, where

$$\ker\big(p(\sigma)\big) := \{\, y \mid p_0 y + p_1 \sigma y + \cdots + p_{\mathtt{n}} \sigma^{\mathtt{n}} y = 0 \,\}. \tag{8}$$

Then, the kernel representation $\ker\big(p(\sigma)\big)$ of the common dynamics $\mathscr{B}$ is given by the greatest common divisor $p = \gcd(p^1, \cdots, p^N)$. In the case when $\mathscr{B}_1, \ldots, \mathscr{B}_N$ have no common dynamics ($\mathscr{B} = \{\, 0 \,\}$), a problem of finding approximate common dynamics of a specified dimension is considered. The approximate common dynamics problem is equivalent to the approximate common divisor Problem (1).

A variation of the common dynamic's estimation problem is considered in [11]. In this case, which we call "data-driven" in order to distinguish it from the "model-based" Problem 2, the aiming is to model a set of scalar time series $y^1, \ldots, y^N$ by sums of, respectively, $\mathtt{n}_1, \ldots, \mathtt{n}_N$ damped exponentials, which have $\mathtt{n}_{\mathrm{c}} \leq \min(\mathtt{n}_1, \ldots, \mathtt{n}_N)$ common exponents. The given time series

$$y^i = \big(y^i(1), \ldots, y^i(T_i)\big)$$

are approximated by time series $\widehat{y}$ satisfying the model equation

$$\widehat{y}^i(t) = \sum_{j=1}^{\mathtt{n}_i - \mathtt{n}_{\mathrm{c}}} \alpha_{ij} \lambda_{ij}^t + \sum_{j=1}^{\mathtt{n}_{\mathrm{c}}} \beta_{ij} \mu_j^t, \qquad t = 1, \ldots, T_i. \tag{9}$$

Here, $\mu_1, \ldots, \mu_{\mathtt{n}_{\mathrm{c}}}$ are the exponents common to all signals and $\lambda_{i1}, \ldots, \lambda_{i\mathtt{n}_i}$ are the remaining exponents of the ith signal.

In [11], a subspace-type method for common dynamics estimation is proposed. Assuming that the data is generated in the output error setup, *i.e.*, $y^i = \bar{y}^i + \tilde{y}^i$, where the true values $\bar{y}^i$ satisfy the model (9) and $\tilde{y}^i$ is the measurement noise that is zero mean, white, Gaussian, the maximum-likelihood estimator is

$$\begin{aligned} &\text{minimize} \quad \text{over } \widehat{y}^i \in \mathbb{R}^{T_i}, \lambda_{ij} \in \mathbb{C}, \mu_j \in \mathbb{C}, \alpha_{ij} \in \mathbb{C}, \text{ and } \beta_{ij} \in \mathbb{C} \quad \sqrt{\sum_{i=1}^{N} \|y^i - \widehat{y}^i\|_2^2} \\ &\text{subject to} \quad (9). \end{aligned} \tag{10}$$

The following result shows equivalent optimization problems to (10) based on the kernel and state space representations of the model.

Theorem 5. *Problem* (10) *is equivalent to the following problems:*

– *kernel representation*

$$\begin{aligned} &\text{minimize} \quad \textit{over } \widehat{y}^i \in \mathbb{R}^{T_i}, R_{s,i}, R_c \quad \|y - \widehat{y}\|_2 \\ &\textit{subject to} \quad (R_{s,i} \star R_c) \mathscr{H}_{\mathtt{n}_i+1}(\widehat{y}^i) = 0, \quad \textit{for } i = 1, \ldots, N. \end{aligned} \tag{11}$$

– *state-space representation*

$$\begin{aligned} &\text{minimize} \quad \textit{over } \widehat{y}^i \in \mathbb{R}^{T_i}, \lambda_i, \mu, c_i, c' \quad \|y - \widehat{y}\|_2 \\ &\textit{subject to} \quad \widehat{y} \in \mathscr{B}\Big(\operatorname{diag}(\lambda_1, \ldots, \lambda_N, \mu), \begin{bmatrix} c_1 & & c_1' \\ & \ddots & \vdots \\ & & c_N \; c_N' \end{bmatrix} \Big), \end{aligned} \tag{12}$$

where $\lambda_i \in \mathbb{C}^{1\times(\ell_i-\ell_c)}$, $c_i \in \mathbb{C}^{1\times(\ell_i-\ell_c)}$, *and* $c' \in \mathbb{C}^{1\times\ell_c}$.

Although these problems are not equivalent to the approximate common divisor problem (1), the solution methods are closely related. Indeed (11) is a Hankel structured low-rank approximation problem. As shown in [15], Problem (1) is a Sylvester structured low-rank approximation problem.

Acknowledgements. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant 258581 "Structured low-rank approximation: Theory, algorithms, and applications"; Fund for Scientific Research (FWO-Vlaanderen), FWO projects G028015N "Decoupling multivariate polynomials in nonlinear system identification"; G090117N "Block-oriented nonlinear identification using Volterra series"; and FWO/FNRS Excellence of Science project 30468160 "Structured low-rank matrix/tensor approximation: numerical optimization-based algorithms and applications".

References

1. Bini, D.A., Boito, P.: A fast algorithm for approximate polynomial GCD based on structured matrix computations. In: Bini, D.A., Mehrmann, V., Olshevsky, V., Tyrtyshnikov, E.E., van Barel, M. (eds.) Numerical Methods for Structured Matrices and Applications. Operator Theory: Advances and Applications, vol. 199, pp. 155–173. Birkhäuser, Basel (2010). https://doi.org/10.1007/978-3-7643-8996-3_6
2. Eising, R.: Distance between controllable and uncontrollable. Control Lett. **4**, 263–264 (1984)
3. Gu, M., Mengi, E., Overton, M., Xia, J., Zhu, J.: Fast methods for estimating the distance to uncontrollability. SIAM J. Matrix Anal. Appl. **28**, 477–502 (2006)
4. Higham, N.: Matrix nearness problems and applications. In: Gover, M., Barnett, S. (eds.) Applications of Matrix Theory, pp. 1–27. Oxford University Press, Oxford (1989)
5. Hu, G., Davison, E.: Real controllablity/stabilizability radius of LTI systems. IEEE Trans. Automat. Control **49**, 254–258 (2004)
6. Kaltofen, E., Corless, R.M., Jeffrey, D.J.: Challenges of symbolic computation: my favorite open problems. J. Symbolic Comput. **29**(6), 891–919 (2000)
7. Karow, M., Kressner, D.: On the structured distance to uncontrollability. Control Lett. **58**, 128–132 (2009)
8. Khare, S., Pillai, H., Belur, M.: Computing the radius of controllability for state space systems. Control Lett. **61**, 327–333 (2012)
9. Markovsky, I., Van Huffel, S.: An algorithm for approximate common divisor computation. In: Proceedings of the 17th Symposium on Mathematical Theory of Networks and Systems, Kyoto, Japan, pp. 274–279 (2006)
10. Paige, C.C.: Properties of numerical algorithms related to computing controllability. IEEE Trans. Automat. Control **26**, 130–138 (1981)
11. Papy, J.M., Lathauwer, L.D., Huffel, S.V.: Common pole estimation in multi-channel exponential data modeling. Signal Process. **86**(4), 846–858 (2006)
12. Polderman, J., Willems, J.C.: Introduction to Mathematical Systems Theory. Springer, New York (1998). https://doi.org/10.1007/978-1-4757-2953-5
13. Qiu, W., Hua, Y., Abed-Meraim, K.: A subspace method for the computation of the GCD of polynomials. Automatica **33**(4), 741–743 (1997)
14. Rupprecht, D.: An algorithm for computing certified approximate GCD of n univariate polynomials. J. Pure Appl. Algebra **139**(1–3), 255–284 (1999)
15. Usevich, K., Markovsky, I.: Variable projection methods for approximate (greatest) common divisor computations. Theor. Comput. Sci. **681**, 176–198 (2017)
16. Willems, J.C.: Paradigms and puzzles in the theory of dynamical systems. IEEE Trans. Automat. Control **36**(3), 259–294 (1991)
17. Zeng, Z., Dayton, B.: The approximate GCD of inexact polynomials. Part I: a univariate algorithm. In: Proceedings of the International Symposium on Symbolic and Algebraic Computation, pp. 320–327 (2004)
18. Zhi, L., Yang, Z.: Computing approximate GCD of univariate polynomials by structure total least norm. In: Wang, D., Zhi, L. (eds.) International Workshop on Symbolic-Numeric, Xian, China, pp. 188–201 (2005)